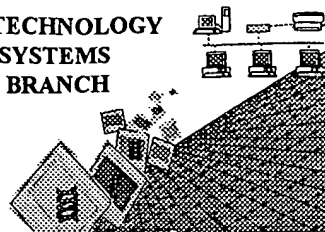


BIOTECHNOLOGY
SYSTEMS
BRANCH



DL

SUB
1231-2001

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/674,368 A
Source: Per/09
Date Processed by STIC: 2/6/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER** **VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 09/674,368A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) 2,4. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001

BEST AVAILABLE COPY



PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,368A

DATE: 02/06/2002
TIME: 14:15:57

ml-3

Input Set : A:\33377PCT.APP.txt
Output Set: N:\CRF3\02062002\I674368A.raw

**Does Not Comply
Corrected Diskette Needed**

5 <110> APPLICANT: American Cyanamid Company
7 <120> TITLE OF INVENTION: Vaccines Containing Recombinant Pilin Against Neisseria
9 Gonorrhoeae or Neisseria Meningitidis
11 <130> FILE REFERENCE: 33377-00/PCT
13 <140> CURRENT APPLICATION NUMBER: US/09/674,368A
14 <141> CURRENT FILING DATE: 2001-12-11
16 <160> NUMBER OF SEQ ID NOS: 24
18 <170> SOFTWARE: PatentIn Ver. 2.0
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 504
22 <212> TYPE: DNA
23 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Description of Artificial Sequence:Chimeric of
27 Neisseria Meningitidis Class I and Neisseria
28 Gonorrhoeae
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (1)..(501)
34 <400> SEQUENCE: 1
36 atg gat acc ctt caa aaa ggc ttt acc ctt atc gag ctg atg att gtg 48
37 Met Asp Thr Leu Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val
38 1 5 10 15
40 atc gcc atc gtc ggc att ttg gcg gca gtc gcc ctt ccc gcc tac caa 96
41 Ile Ala Ile Val Gly Ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln
42 20 25 30
44 gac tac acc gcc cgc gcg caa gtt tcc gaa gcc atc ctt ttg gcc gaa 144
45 Asp Tyr Thr Ala Arg Ala Gln Val Ser Glu Ala Ile Leu Leu Ala Glu
46 35 40 45
48 ggt caa aaa tca gcc gtt acc gag tat tac ctg aat cac ggc gaa tgg 192
49 Gly Gln Lys Ser Ala Val Thr Glu Tyr Tyr Leu Asn His Gly Glu Trp
50 50 55 60
52 ccc ggc aac aac act tct gcc ggc gtg gca tct tct tca aca atc aaa 240
53 Pro Gly Asn Asn Thr Ser Ala Gly Val Ala Ser Ser Ser Thr Ile Lys
54 65 70 75 80
58 ggc aaa tat gtt aag gaa gtt aca gtc gca aac ggc gtc att acc gcc 288
59 Gly Lys Tyr Val Lys Glu Val Thr Val Ala Asn Gly Val Ile Thr Ala
W--> 60 85 90 95 95 *← fix numbering*
62 aca atg ctt tca agc ggc gta aac aaa gaa atc caa ggc aaa aaa ctc 336
63 Thr Met Leu Ser Ser Gly Val Asn Lys Glu Ile Gln Gly Lys Lys Leu
64 100 105 110
66 tcc ctg tgg gcc aag cgt caa gac ggt tgc gta aaa tgg ttc tgc gga 384
67 Ser Leu Trp Ala Lys Arg Gln Asp Gly Ser Val Lys Trp Phe Cys Gly

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/674,368A

TIME: 14:15:57

Input Set : A:\33377PCT.APP.txt

Output Set: N:\CRF3\02062002\I674368A.raw

W--> 68 115 120 125
 70 cag ccg gtt acg cgc acc gac gcc aaa gcc gac acc gtc gcc gcc gcc 432
 71 Gln Pro Val Thr Arg Thr Asp Ala Lys Ala Asp Thr Val Ala Ala Ala
 W--> 72 130 135 140
 74 gcc aag acc gcc gac aac atc aac acc aag cac ctg ccg tca acc tgc 480
 75 Ala Lys Thr Ala Asp Asn Ile Asn Thr Lys His Leu Pro Ser Thr Cys
 W--> 76 145 150 155 160
 78 cgc gac gca agt gat gcc agc taa 504
 79 Arg Asp Ala Ser Asp Ala Ser

W--> 80 165
 82 <210> SEQ ID NO: 2
 83 <211> LENGTH: 167
 84 <212> TYPE: PRT
 85 <213> ORGANISM: Artificial Sequence
 W--> 87 <220> FEATURE:
 W--> 87 <223> OTHER INFORMATION:
 87 <400> SEQUENCE: 2
 88 Met Asp Thr Leu Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val
 89 1 5 10 15
 91 Ile Ala Ile Val Gly Ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln
 92 20 25 30
 95 Asp Tyr Thr Ala Arg Ala Gln Val Ser Glu Ala Ile Leu Leu Ala Glu
 96 35 40 45
 99 Gly Gln Lys Ser Ala Val Thr Glu Tyr Tyr Leu Asn His Gly Glu Trp
 100 50 55 60
 103 Pro Gly Asn Asn Thr Ser Ala Gly Val Ala Ser Ser Ser Thr Ile Lys
 104 65 70 75 80
 107 Gly Lys Tyr Val Lys Glu Val Thr Val Ala Asn Gly Val Ile Thr Ala
 108 85 90 95
 111 Thr Met Leu Ser Ser Gly Val Asn Lys Glu Ile Gln Gly Lys Lys Leu
 112 100 105 110
 115 Ser Leu Trp Ala Lys Arg Gln Asp Gly Ser Val Lys Trp Phe Cys Gly
 116 115 120 125
 118 Gln Pro Val Thr Arg Thr Asp Ala Lys Ala Asp Thr Val Ala Ala Ala
 119 130 135 140
 121 Ala Lys Thr Ala Asp Asn Ile Asn Thr Lys His Leu Pro Ser Thr Cys
 122 145 150 155 160
 124 Arg Asp Ala Ser Asp Ala Ser
 125 165
 128 <210> SEQ ID NO: 3
 129 <211> LENGTH: 510
 130 <212> TYPE: DNA
 131 <213> ORGANISM: Artificial Sequence
 133 <220> FEATURE:
 134 <223> OTHER INFORMATION: Description of Artificial Sequence:Chimeric of
 135 Neisseria Meningitidis Class II and Neisseria
 136 Gonorrhoeae
 138 <220> FEATURE:
 139 <221> NAME/KEY: CDS

see item 6 on Error Summary Sheet

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/674,368A

TIME: 14:15:57

Input Set : A:\33377PCT.APP.txt

Output Set: N:\CRF3\02062002\I674368A.raw

```

140 <222> LOCATION: (1)..(510)
142 <400> SEQUENCE: 3
143 atg gaa gca atc caa aaa ggt ttc acc ctg atc gag ctg atg atc gtc 48
144 Met Glu Ala Ile Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val
145 1 5 10 15
147 atc gcc atc gtc ggt atc ttg gca gcc gtc gcc ctg ccc gca tac caa 96
148 Ile Ala Ile Val Gly Ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln
149 20 25 30
151 gac tac acc gcg cgc gcc caa atg tcc gaa gcc ctg act ttg gca gaa 144
152 Asp Tyr Thr Ala Arg Ala Gln Met Ser Glu Ala Leu Thr Leu Ala Glu
153 35 40 45
155 ggt caa aaa tcc gca gtg atc gag tat tat tcc gac aac ggc aca ttc 192
156 Gly Gln Lys Ser Ala Val Ile Glu Tyr Tyr Ser Asp Asn Gly Thr Phe
157 50 55 60
159 ccg aac agc aat act tcc gca ggt att gct gcc tct aac gag att aaa 240
160 Pro Asn Ser Asn Thr Ser Ala Gly Ile Ala Ala Ser Asn Glu Ile Lys
161 65 70 75 80
163 ggt aag tat gtg gca tcg gtt aag gtt gaa ggt aat gcc tct gtt gct 288
164 Gly Lys Tyr Val Ala Ser Val Lys Val Glu Gly Asn Ala Ser Val Ala
165 85 90 95
167 tct att acc gct acc atg aac tct agt aat gtg aat aag gac atc aaa 336
168 Ser Ile Thr Ala Thr Met Asn Ser Ser Asn Val Asn Lys Asp Ile Lys
169 100 105 110
172 ggt aaa acc ttg gta ctc gtc ggc aaa caa aac tcc ggt tcg gta aaa 384
173 Gly Lys Thr Leu Val Leu Val Gly Lys Gln Asn Ser Gly Ser Val Lys
174 115 120 125
176 tgg ttc tgc gga cag ccg gtt acg cgc gac aac gcc gac aac gac gac 432
177 Trp Phe Cys Gly Gln Pro Val Thr Arg Asp Asn Ala Asp Asn Asp Asp
178 130 135 140
180 gtc aaa gac gcc ggc aac aac ggc atc gaa acc aag cac ctg ccg tca 480
181 Val Lys Asp Ala Gly Asn Asn Gly Ile Glu Thr Lys His Leu Pro Ser
182 145 150 155 160
184 acc tgc cgc gat acg tca tct gat gcc aaa 510
185 Thr Cys Arg Asp Thr Ser Ser Asp Ala Lys
186 165 170

```

189 <210> SEQ ID NO: 4

190 <211> LENGTH: 170

191 <212> TYPE: PRT

192 <213> ORGANISM: Artificial Sequence

W--> 194 <220> FEATURE:

W--> 194 <223> OTHER INFORMATION:

194 <400> SEQUENCE: 4

```

195 Met Glu Ala Ile Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val
196 1 5 10 15
198 Ile Ala Ile Val Gly Ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln
199 20 25 30
201 Asp Tyr Thr Ala Arg Ala Gln Met Ser Glu Ala Leu Thr Leu Ala Glu
202 35 40 45
204 Gly Gln Lys Ser Ala Val Ile Glu Tyr Tyr Ser Asp Asn Gly Thr Phe

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/674,368A

TIME: 14:15:57

Input Set : A:\33377PCT.APP.txt

Output Set: N:\CRF3\02062002\I674368A.raw

```

205      50      55      60
207 Pro Asn Ser Asn Thr Ser Ala Gly Ile Ala Ala Ser Asn Glu Ile Lys
208 65      70      75      80
210 Gly Lys Tyr Val Ala Ser Val Lys Val Glu Gly Asn Ala Ser Val Ala
211      85      90      95
213 Ser Ile Thr Ala Thr Met Asn Ser Ser Asn Val Asn Lys Asp Ile Lys
214      100      105      110
216 Gly Lys Thr Leu Val Leu Val Gly Lys Gln Asn Ser Gly Ser Val Lys
217      115      120      125
219 Trp Phe Cys Gly Gln Pro Val Thr Arg Asp Asn Ala Asp Asn Asp Asp
220      130      135      140
222 Val Lys Asp Ala Gly Asn Asn Gly Ile Glu Thr Lys His Leu Pro Ser
223 145      150      155      160
225 Thr Cys Arg Asp Thr Ser Ser Asp Ala Lys
226      165      170
230 <210> SEQ ID NO: 5
231 <211> LENGTH: 30
232 <212> TYPE: DNA
233 <213> ORGANISM: Neisseria gonorrhoeae
235 <400> SEQUENCE: 5
236 ccccgcgcca tggataccct tcaaaaaggc 30
239 <210> SEQ ID NO: 6
240 <211> LENGTH: 30
241 <212> TYPE: DNA
242 <213> ORGANISM: Neisseria gonorrhoeae
244 <400> SEQUENCE: 6
245 gggcctggat ccgtgggaaa tcacttaccg 30
248 <210> SEQ ID NO: 7
249 <211> LENGTH: 29
250 <212> TYPE: DNA
251 <213> ORGANISM: Neisseria gonorrhoeae
253 <400> SEQUENCE: 7
254 ggctctagac tgtcagacca agtttactc 29
256 <210> SEQ ID NO: 8
257 <211> LENGTH: 27
258 <212> TYPE: DNA
259 <213> ORGANISM: Neisseria gonorrhoeae
261 <400> SEQUENCE: 8
262 ggctctagat tgaagcattt atcaggg 27
264 <210> SEQ ID NO: 9
265 <211> LENGTH: 28
266 <212> TYPE: DNA
267 <213> ORGANISM: Neisseria gonorrhoeae
269 <400> SEQUENCE: 9
270 ggctctagat aaacagtaat acaagggg 28
272 <210> SEQ ID NO: 10
273 <211> LENGTH: 28
274 <212> TYPE: DNA
275 <213> ORGANISM: Neisseria gonorrhoeae

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/674,368A

TIME: 14:15:57

Input Set : A:\33377PCT.APP.txt

Output Set: N:\CRF3\02062002\I674368A.raw

```

277 <400> SEQUENCE: 10
278 ggctctagat tagaaaaact catcgagc 28
281 <210> SEQ ID NO: 11
282 <211> LENGTH: 36
283 <212> TYPE: DNA
284 <213> ORGANISM: Neisseria Meningitidis Class I
287 <400> SEQUENCE: 11
288 ccccgcgcca tggacaccct tcaaaaaggt ttacc 36
290 <210> SEQ ID NO: 12
291 <211> LENGTH: 39
292 <212> TYPE: DNA
293 <213> ORGANISM: Neisseria Meningitidis Class I
295 <400> SEQUENCE: 12
296 gggcctggat ccgagtggcc gtggaaaatc acttaccgc 39
298 <210> SEQ ID NO: 13
299 <211> LENGTH: 31
300 <212> TYPE: DNA
302 <213> ORGANISM: Neisseria Meningitidis Class I
304 <400> SEQUENCE: 13
305 ccggcgcgtc tctcacggcg aatggcccgc c 31
307 <210> SEQ ID NO: 14
308 <211> LENGTH: 39
309 <212> TYPE: DNA
310 <213> ORGANISM: Neisseria Meningitidis Class I
312 <400> SEQUENCE: 14
313 gggcctggat ccgagtggcc gtggaaaatc acttaccgc 39
316 <210> SEQ ID NO: 15
317 <211> LENGTH: 25
318 <212> TYPE: DNA
319 <213> ORGANISM: Neisseria gonorrhoeae
321 <400> SEQUENCE: 15
322 gcataattcg tgtcgctcaa ggcgc 25
324 <210> SEQ ID NO: 16
325 <211> LENGTH: 34
326 <212> TYPE: DNA
327 <213> ORGANISM: Neisseria gonorrhoeae
329 <400> SEQUENCE: 16
330 gccgcgcgtc tcccgatgatt caggtaatac tcgg 34
332 <210> SEQ ID NO: 17
333 <211> LENGTH: 25
334 <212> TYPE: DNA
335 <213> ORGANISM: Neisseria Meningitidis Class I
337 <400> SEQUENCE: 17
338 gcataattcg tgtcgctcaa ggcgc 25
343 <210> SEQ ID NO: 18
344 <211> LENGTH: 39
345 <212> TYPE: DNA
346 <213> ORGANISM: Neisseria Meningitidis Class I
348 <400> SEQUENCE: 18

```

VERIFICATION SUMMARY

DATE: 02/06/2002

PATENT APPLICATION: US/09/674,368A

TIME: 14:15:59

Input Set : A:\33377PCT.APP.txt

Output Set: N:\CRF3\02062002\I674368A.raw

L:13 M:270 C: Current Application Number differs, Replaced Application Number
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:60 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:64 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:68 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:72 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:76 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:80 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:87 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:87 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:194 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:194 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION: